



SECOND SUBSTITUTE SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Pulst, Stefan M

(ii) TITLE OF INVENTION: NUCLEIC ACID ENCODING SPINOCEREBELLAR ATAXIA-2 AND PRODUCTS RELATED THERETO

(iii) NUMBER OF SEQUENCES: 19

(iv) CORRESPONDENCE ADDRESS:

- (A) ADDRESSEE: Mueting, Raasch & Gebhardt, P.A.
- (B) STREET: 119 North Fourth Street
- (C) CITY: Minneapolis
- (D) STATE: Minnesota
- (E) COUNTRY: USA
- (F) ZIP: 55401

(v) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER: US 08/727,084
- (B) FILING DATE: 08-OCT-1996
- (C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: Mueting, Ann M.
- (B) REGISTRATION NUMBER: 33,977
- (C) REFERENCE/DOCKET NUMBER: 232.00010101

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: 612/305-1220
- (B) TELEFAX: 612/305-1228

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 516 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: both

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TTGGTAGCAA CGGAAACGGC GGCGGCGCGT TTGGCCCGG CTCCGGCGG CTCCGGTC	60
TCGGCGGGCC TCCCCGCCCC TTGTCGTCG TCCTTCTCCC CCTCGCCAGC CGGGCGCCC	120
CTCCGGCCGC GCCAACCCGC GCCTCCCCGC TCGGCGCCCG TGCGTCCCCG CGCGTTCCG	180
GCGTCTCCTT GGCGCGCCCG GCTCCGGCT GTCCCCGCC GGCGTGCGAG CGGGTGTATG	240
GGCCCCTCAC CATGTCGCTG AAGCCCCAGC AGCAGCAGCA GCAGCAGCAG CAACAGCAGC	300
AGCAGCAACA GCAGCAGCAG CAGCAGCAGC AGCCGCCGCC CGCGGCTGCC AATGTCCGCA	360
AGCCCGGCCGG CAGCGGCCTT CTAGCGTCGC CCGCCGCCGC GCCTCGCCG TCCTCGTCCT	420
CGGTCTCCTC GTCTCGGCC ACGGCTCCCT CCTCGGTGGT CGCGGCGACC TCCGGCGGCCG	480
GGAGGCCCCGG CCTGGGCAGG TGGGTGTCCG CACCCCC	516

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4481 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: both

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 163..4101

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

ACCCCGAGA AAGCAACCCA GCGCGCCGCC CGCTCCTCAC GTGTCCCTCC CGGCCCCGGG	60
GCCACCTCAC GTTCTGCTTC CGTCTGACCC CTCCGACTTC CGGTAAAGAG TCCCTATCCG	120
CACCTCCGCT CCCACCCGGC GCCTCGGCCGC GCCCGCCCTC CG ATG CGC TCA GCG Met Arg Ser Ala 1	174
GCC GCA GCT CCT CGG AGT CCC GCG GTG GCC ACC GAG TCT CGC CGC TTC Ala Ala Ala Pro Arg Ser Pro Ala Val Ala Thr Glu Ser Arg Arg Phe	222
5 10 15 20	

GCC GCA GCC AGG TGG CCC GGG TGG CGC TCG CTC CAG CGG CCG GCG CGG Ala Ala Ala Arg Trp Pro Gly Trp Arg Ser Leu Gln Arg Pro Ala Arg	25	30	35	270
CGG AGC GGG CGG GGC GGC GGT GGC GCG GCC CCG GGA CCG TAT CCC TCC Arg Ser Gly Arg Gly Gly Ala Ala Pro Gly Pro Tyr Pro Ser	40	45	50	318
GCC GCC CCT CCC CCG CCC GGC CCC GGC CCC CCT CCC TCC CGG CAG AGC Ala Ala Pro Pro Pro Pro Gly Pro Gly Pro Pro Ser Arg Gln Ser	55	60	65	366
TCG CCT CCC TCC GCC TCA GAC TGT TTT GGT AGC AAC GGC AAC GGC GGC Ser Pro Pro Ser Ala Ser Asp Cys Phe Gly Ser Asn Gly Asn Gly Gly	70	75	80	414
GGC GCG TTT CGG CCC GGC TCC CGG CGG CTC CTT GGT CTC GGC GGG CCT Gly Ala Phe Arg Pro Gly Ser Arg Arg Leu Leu Gly Leu Gly Pro	85	90	95	462
CCC CGC CCC TTC GTC GTC GTC CTT CTC CCC CTC GCC AGC CCG GGC GCC Pro Arg Pro Phe Val Val Val Leu Leu Pro Leu Ala Ser Pro Gly Ala	105	110	115	510
CCT CCG GCC GCG CCA ACC CGC GCC TCC CCG CTC GGC GCC CGT GCG TCC Pro Pro Ala Ala Pro Thr Arg Ala Ser Pro Leu Gly Ala Arg Ala Ser	120	125	130	558
CCG CCG CGT TCC GGC GTC TCC TTG GCG CGC CCG GCT CCC GGC TGT CCC Pro Pro Arg Ser Gly Val Ser Leu Ala Arg Pro Ala Pro Gly Cys Pro	135	140	145	606
CGC CCG GCG TGC GAG CCG GTG TAT GGG CCC CTC ACC ATG TCG CTG AAG Arg Pro Ala Cys Glu Pro Val Tyr Gly Pro Leu Thr Met Ser Leu Lys	150	155	160	654
CCC CAG CAG CAG CAG CAG CAG CAG CAA CAG CAG CAG CAG CAA CAG Pro Gln	165	170	175	702
CAG CAG CAG CAG CAG CAG CCG CCG CCC GCG GCT GCC AAT GTC CGC Gln Gln Gln Gln Gln Pro Pro Ala Ala Ala Asn Val Arg	185	190	195	750
AAG CCC GGC AGC GGC CTT CTA GCG TCG CCC GCC GCC GCG CCT TCG Lys Pro Gly Gly Ser Gly Leu Leu Ala Ser Pro Ala Ala Pro Ser	200	205	210	798
CCG TCC TCG TCC TCG GTC TCC TCG TCG GCC ACG GCT CCC TCC TCG Pro Ser Ser Ser Ser Val Ser Ser Ser Ala Thr Ala Pro Ser Ser	215	220	225	846

GTG GTC GCG GCG ACC TCC GGC GGC GGG AGG CCC GGC CTG GGC AGA GGT		894
Val Val Ala Ala Thr Ser Gly Gly Gly Arg Pro Gly Leu Gly Arg Gly		
230	235	240
CGA AAC AGT AAC AAA GGA CTG CCT CAG TCT ACG ATT TCT TTT GAT GGA		942
Arg Asn Ser Asn Lys Gly Leu Pro Gln Ser Thr Ile Ser Phe Asp Gly		
245	250	255
260		
ATC TAT GCA AAT ATG AGG ATG GTT CAT ATA CTT ACA TCA GTT GTT GGC		990
Ile Tyr Ala Asn Met Arg Met Val His Ile Leu Thr Ser Val Val Gly		
265	270	275
TCC AAA TGT GAA GTA CAA GTG AAA AAT GGA GGT ATA TAT GAA GGA GTT		1038
Ser Lys Cys Glu Val Gln Val Lys Asn Gly Gly Ile Tyr Glu Gly Val		
280	285	290
TTT AAA ACT TAC AGT CCG AAG TGT GAT TTG GTA CTT GAT GCC GCA CAT		1086
Phe Lys Thr Tyr Ser Pro Lys Cys Asp Leu Val Leu Asp Ala Ala His		
295	300	305
GAG AAA AGT ACA GAA TCC AGT TCG GGG CCG AAA CGT GAA GAA ATA ATG		1134
Glu Lys Ser Thr Glu Ser Ser Gly Pro Lys Arg Glu Glu Ile Met		
310	315	320
GAG AGT ATT TTG TTC AAA TGT TCA GAC TTT GTT GTG GTA CAG TTT AAA		1182
Glu Ser Ile Leu Phe Lys Cys Ser Asp Phe Val Val Val Gln Phe Lys		
325	330	335
340		
GAT ATG GAC TCC AGT TAT GCA AAA AGA GAT GCT TTT ACT GAC TCT GCT		1230
Asp Met Asp Ser Ser Tyr Ala Lys Arg Asp Ala Phe Thr Asp Ser Ala		
345	350	355
ATC AGT GCT AAA GTG AAT GGC GAA CAC AAA GAG AAG GAC CTG GAG CCC		1278
Ile Ser Ala Lys Val Asn Gly Glu His Lys Glu Lys Asp Leu Glu Pro		
360	365	370
TGG GAT GCA GGT GAA CTC ACA GCC AAT GAG GAA CTT GAG GCT TTG GAA		1326
Trp Asp Ala Gly Glu Leu Thr Ala Asn Glu Glu Leu Glu Ala Leu Glu		
375	380	385
AAT GAC GTA TCT AAT GGA TGG GAT CCC AAT GAT ATG TTT CGA TAT AAT		1374
Asn Asp Val Ser Asn Gly Trp Asp Pro Asn Asp Met Phe Arg Tyr Asn		
390	395	400
GAA GAA AAT TAT GGT GTA GTG TCT ACG TAT GAT AGC AGT TTA TCT TCG		1422
Glu Glu Asn Tyr Gly Val Val Ser Thr Tyr Asp Ser Ser Leu Ser Ser		
405	410	415
420		
TAT ACA GTG CCC TTA GAA AGA GAT AAC TCA GAA GAA TTT TTA AAA CGG		1470
Tyr Thr Val Pro Leu Glu Arg Asp Asn Ser Glu Glu Phe Leu Lys Arg		
425	430	435

GAA GCA AGG GCA AAC CAG TTA GCA GAA GAA ATT GAG TCA AGT GCC CAG Glu Ala Arg Ala Asn Gln Leu Ala Glu Glu Ile Glu Ser Ser Ala Gln 440 445 450	1518
TAC AAA GCT CGA GTG GCC CTG GAA AAT GAT GAT AGG AGT GAG GAA GAA Tyr Lys Ala Arg Val Ala Leu Glu Asn Asp Asp Arg Ser Glu Glu Glu 455 460 465	1566
AAA TAC ACA GCA GTT CAG AGA AAT TCC AGT GAA CGT GAG GGG CAC AGC Lys Tyr Thr Ala Val Gln Arg Asn Ser Ser Glu Arg Glu Gly His Ser 470 475 480	1614
ATA AAC ACT AGG GAA AAT AAA TAT ATT CCT CCT GGA CAA AGA AAT AGA Ile Asn Thr Arg Glu Asn Lys Tyr Ile Pro Pro Gly Gln Arg Asn Arg 485 490 495 500	1662
GAA GTC ATA TCC TGG GGA AGT GGG AGA CAG AAT TCA CCG CGT ATG GGC Glu Val Ile Ser Trp Gly Ser Gly Arg Gln Asn Ser Pro Arg Met Gly 505 510 515	1710
CAG CCT GGA TCG GGC TCC ATG CCA TCA AGA TCC ACT TCT CAC ACT TCA Gln Pro Gly Ser Gly Ser Met Pro Ser Arg Ser Thr Ser His Thr Ser 520 525 530	1758
GAT TTC AAC CCG AAT TCT GGT TCA GAC CAA AGA GTA GTT AAT GGA GGT Asp Phe Asn Pro Asn Ser Gly Ser Asp Gln Arg Val Val Asn Gly Gly 535 540 545	1806
GTT CCC TGG CCA TCG CCT TGC CCA TCT CCT TCC TCT CGC CCA CCT TCT Val Pro Trp Pro Ser Pro Cys Pro Ser Pro Ser Arg Pro Pro Ser 550 555 560	1854
CGC TAC CAG TCA GGT CCC AAC TCT CTT CCA CCT CGG GCA GCC ACC CCT Arg Tyr Gln Ser Gly Pro Asn Ser Leu Pro Pro Arg Ala Ala Thr Pro 565 570 575 580	1902
ACA CGG CCG CCC TCC AGG CCC CCC TCG CGG CCA TCC AGA CCC CCG TCT Thr Arg Pro Pro Ser Arg Pro Pro Ser Arg Pro Ser Arg Pro Pro Ser 585 590 595	1950
CAC CCC TCT GCT CAT GGT TCT CCA GCT CCT GTC TCT ACT ATG CCT AAA His Pro Ser Ala His Gly Ser Pro Ala Pro Val Ser Thr Met Pro Lys 600 605 610	1998
CGC ATG TCT TCA GAA GGG CCT CCA AGG ATG TCC CCA AAG GCC CAG CGA Arg Met Ser Ser Glu Gly Pro Pro Arg Met Ser Pro Lys Ala Gln Arg 615 620 625	2046
CAT CCT CGA AAT CAC AGA GTT TCT GCT GGG AGG GGT TCC ATA TCC AGT His Pro Arg Asn His Arg Val Ser Ala Gly Arg Gly Ser Ile Ser Ser 630 635 640	2094

GGC CTA GAA TTT GTA TCC CAC AAC CCA CCC AGT GAA GCA GCT ACT CCT Gly Leu Glu Phe Val Ser His Asn Pro Pro Ser Glu Ala Ala Thr Pro 645 650 655 660	2142
CCA GTA GCA AGG ACC AGT CCC TCG GGG GGA ACG TGG TCA TCA GTG GTC Pro Val Ala Arg Thr Ser Pro Ser Gly Gly Thr Trp Ser Ser Val Val 665 670 675	2190
AGT GGG GTT CCA AGA TTA TCC CCT AAA ACT CAT AGA CCC AGG TCT CCC Ser Gly Val Pro Arg Leu Ser Pro Lys Thr His Arg Pro Arg Ser Pro 680 685 690	2238
AGA CAG AAC AGT ATT GGA AAT ACC CCC AGT GGG CCA GTT CTT GCT TCT Arg Gln Asn Ser Ile Gly Asn Thr Pro Ser Gly Pro Val Leu Ala Ser 695 700 705	2286
CCC CAA GCT GGT ATT ATT CCA ACT GAA GCT GTT GCC ATG CCT ATT CCA Pro Gln Ala Gly Ile Ile Pro Thr Glu Ala Val Ala Met Pro Ile Pro 710 715 720	2334
GCT GCA TCT CCT ACG CCT GCT AGT CCT GCA TCG AAC AGA GCT GTT ACC Ala Ala Ser Pro Thr Pro Ala Ser Pro Ala Ser Asn Arg Ala Val Thr 725 730 735 740	2382
CCT TCT AGT GAG GCT AAA GAT TCC AGG CTT CAA GAT CAG AGG CAG AAC Pro Ser Ser Glu Ala Lys Asp Ser Arg Leu Gln Asp Gln Arg Gln Asn 745 750 755	2430
TCT CCT GCA GGG AAT AAA GAA AAT ATT AAA CCC AAT GAA ACA TCA CCT Ser Pro Ala Gly Asn Lys Glu Asn Ile Lys Pro Asn Glu Thr Ser Pro 760 765 770	2478
AGC TTC TCA AAA GCT GAA AAC AAA GGT ATA TCA CCA GTT GTT TCT GAA Ser Phe Ser Lys Ala Glu Asn Lys Gly Ile Ser Pro Val Val Ser Glu 775 780 785	2526
CAT AGA AAA CAG ATT GAT GAT TTA AAG AAA TTT AAG AAT GAT TTT AGG His Arg Lys Gln Ile Asp Asp Leu Lys Lys Phe Lys Asn Asp Phe Arg 790 795 800	2574
TTA CAG CCA AGT TCT ACT TCT GAA TCT ATG GAT CAA CTA CTA AAC AAA Leu Gln Pro Ser Ser Thr Ser Glu Ser Met Asp Gln Leu Leu Asn Lys 805 810 815 820	2622
AAT AGA GAG GGA GAA AAA TCA AGA GAT TTG ATC AAA GAC AAA ATT GAA Asn Arg Glu Gly Glu Lys Ser Arg Asp Leu Ile Lys Asp Lys Ile Glu 825 830 835	2670
CCA AGT GCT AAG GAT TCT TTC ATT GAA AAT AGC AGC AGC AAC TGT ACC Pro Ser Ala Lys Asp Ser Phe Ile Glu Asn Ser Ser Asn Cys Thr 840 845 850	2718

AGT GGC AGC AGC AAG CCG AAT AGC CCC AGC ATT TCC CCT TCA ATA CTT Ser Gly Ser Ser Lys Pro Asn Ser Pro Ser Ile Ser Pro Ser Ile Leu 855 860 865	2766
AGT AAC ACG GAG CAC AAG AGG GGA CCT GAG GTC ACT TCC CAA GGG GTT Ser Asn Thr Glu His Lys Arg Gly Pro Glu Val Thr Ser Gln Gly Val 870 875 880	2814
CAG ACT TCC AGC CCA GCA TGT AAA CAA GAG AAA GAC GAT AAG GAA GAG Gln Thr Ser Ser Pro Ala Cys Lys Gln Glu Lys Asp Asp Lys Glu Glu 885 890 895 900	2862
AAG AAA GAC GCA GCT GAG CAA GTT AGG AAA TCA ACA TTG AAT CCC AAT Lys Lys Asp Ala Ala Glu Gln Val Arg Lys Ser Thr Leu Asn Pro Asn 905 910 915	2910
GCA AAG GAG TTC AAC CCA CGT TCC TTC TCT CAG CCA AAG CCT TCT ACT Ala Lys Glu Phe Asn Pro Arg Ser Phe Ser Gln Pro Lys Pro Ser Thr 920 925 930	2958
ACC CCA ACT TCA CCT CGG CCT CAA GCA CAA CCT AGC CCA TCT ATG GTG Thr Pro Thr Ser Pro Arg Pro Gln Ala Gln Pro Ser Pro Ser Met Val 935 940 945	3006
GGT CAT CAA CAG CCA ACT CCA GTT TAT ACT CAG CCT GTT TGT TTT GCA Gly His Gln Gln Pro Thr Pro Val Tyr Thr Gln Pro Val Cys Phe Ala 950 955 960	3054
CCA AAT ATG ATG TAT CCA GTC CCA GTG AGC CCA GGC GTG CAA CCT TTA Pro Asn Met Met Tyr Pro Val Pro Val Ser Pro Gly Val Gln Pro Leu 965 970 975 980	3102
TAC CCA ATA CCT ATG ACG CCC ATG CCA GTG AAT CAA GCC AAG ACA TAT Tyr Pro Ile Pro Met Thr Pro Met Pro Val Asn Gln Ala Lys Thr Tyr 985 990 995	3150
AGA GCA GTA CCA AAT ATG CCC CAA CAG CGG CAA GAC CAG CAT CAT CAG Arg Ala Val Pro Asn Met Pro Gln Gln Arg Gln Asp Gln His His Gln 1000 1005 1010	3198
AGT GCC ATG ATG CAC CCA GCG TCA GCA GCG GGC CCA CCG ATT GCA GCC Ser Ala Met Met His Pro Ala Ser Ala Ala Gly Pro Pro Ile Ala Ala 1015 1020 1025	3246
ACC CCA CCA GCT TAC TCC ACG CAA TAT GTT GCC TAC AGT CCT CAG CAG Thr Pro Pro Ala Tyr Ser Thr Gln Tyr Val Ala Tyr Ser Pro Gln Gln 1030 1035 1040	3294
TTC CCA AAT CAG CCC CTT GTT CAG CAT GTG CCA CAT TAT CAG TCT CAG Phe Pro Asn Gln Pro Leu Val Gln His Val Pro His Tyr Gln Ser Gln 1045 1050 1055 1060	3342

CAT CCT CAT GTC TAT AGT CCT GTA ATA CAG GGT AAT GCT AGA ATG ATG His Pro His Val Tyr Ser Pro Val Ile Gln Gly Asn Ala Arg Met Met 1065 1070 1075	3390
GCA CCA CCA ACA CAC GCC CAG CCT GGT TTA GTA TCT TCT TCA GCA ACT Ala Pro Pro Thr His Ala Gln Pro Gly Leu Val Ser Ser Ala Thr 1080 1085 1090	3438
CAG TAC GGG GCT CAT GAG CAG ACG CAT GCG ATG TAT GCA TGT CCC AAA Gln Tyr Gly Ala His Glu Gln Thr His Ala Met Tyr Ala Cys Pro Lys 1095 1100 1105	3486
TTA CCA TAC AAC AAG GAG ACA AGC CCT TCT TTC TAC TTT GCC ATT TCC Leu Pro Tyr Asn Lys Glu Thr Ser Pro Ser Phe Tyr Phe Ala Ile Ser 1110 1115 1120	3534
ACG GGC TCC CTT GCT CAG CAG TAT GCG CAC CCT AAC GCT ACC CTG CAC Thr Gly Ser Leu Ala Gln Gln Tyr Ala His Pro Asn Ala Thr Leu His 1125 1130 1135 1140	3582
CCA CAT ACT CCA CAC CCT CAG CCT TCA GCT ACC CCC ACT GGA CAG CAG Pro His Thr Pro His Pro Gln Pro Ser Ala Thr Pro Thr Gly Gln Gln 1145 1150 1155	3630
CAA AGC CAA CAT GGT GGA AGT CAT CCT GCA CCC AGT CCT GTT CAG CAC Gln Ser Gln His Gly Ser His Pro Ala Pro Ser Pro Val Gln His 1160 1165 1170	3678
CAT CAG CAC CAG GCC CAG GCT CTC CAT CTG GCC AGT CCA CAG CAG His Gln His Gln Ala Ala Gln Ala Leu His Leu Ala Ser Pro Gln Gln 1175 1180 1185	3726
CAG TCA GCC ATT TAC CAC GCG GGG CTT GCG CCA ACT CCA CCC TCC ATG Gln Ser Ala Ile Tyr His Ala Gly Leu Ala Pro Thr Pro Pro Ser Met 1190 1195 1200	3774
ACA CCT GCC TCC AAC ACG CAG TCG CCA CAG AAT AGT TTC CCA GCA GCA Thr Pro Ala Ser Asn Thr Gln Ser Pro Gln Asn Ser Phe Pro Ala Ala 1205 1210 1215 1220	3822
CAA CAG ACT GTC TTT ACG ATC CAT CCT TCT CAC GTT CAG CCG GCG TAT Gln Gln Thr Val Phe Thr Ile His Pro Ser His Val Gln Pro Ala Tyr 1225 1230 1235	3870
ACC AAC CCA CCC CAC ATG GCC CAC GTA CCT CAG GCT CAT GTA CAG TCA Thr Asn Pro Pro His Met Ala His Val Pro Gln Ala His Val Gln Ser 1240 1245 1250	3918
GGA ATG GTT CCT TCT CAT CCA ACT GCC CAT GCG CCA ATG ATG CTA ATG Gly Met Val Pro Ser His Pro Thr Ala His Ala Pro Met Met Leu Met 1255 1260 1265	3966

ACG ACA CAG CCA CCC GGC GGT CCC CAG GCC GCC CTC GCT CAA AGT GCA	4014
Thr Thr Gln Pro Pro Gly Gly Pro Gln Ala Ala Leu Ala Gln Ser Ala	
1270 1275 1280	
CTA CAG CCC ATT CCA GTC TCG ACA ACA GCG CAT TTC CCC TAT ATG ACG	4062
Leu Gln Pro Ile Pro Val Ser Thr Thr Ala His Phe Pro Tyr Met Thr	
1285 1290 1295 1300	
CAC CCT TCA GTA CAA GCC CAC CAA CAG CAG TTG TAAGGCTGCC	4108
His Pro Ser Val Gln Ala His His Gln Gln Leu	
1305 1310	
CTGGAGGAAC CGAAAGGCCA AATTCCCTCC TCCCTTCTAC TGCTTCTACC AACTGGAAGC	4168
ACAGAAAAC AGAATTTCAT TTATTTGTT TTTAAAATAT ATATGTTGAT TTCTTGTAAC	4228
ATCCAATAGG AATGCTAACCA GTTCACTTGC AGTGGAAAGAT ACTTGGACCG AGTAGAGGCCA	4288
TTTAGGAAC TGGGGCTAT TCCATAATT CATATGCTGT TTCAGAGTCC CGCAGGTACC	4348
CCAGCTCTGC TTGCCGAAAC TGGAAGTTAT TTATTTTTA ATAACCCTTG AAAGTCATGA	4408
ACACATCAGC TAGCAAAAGA AGTAACAAGA GTGATTCTTG CTGCTATTAC TGCTAAAAAA	4468
AAAAAAAAAA AAA	4481

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1312 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met Arg Ser Ala Ala Ala Ala Pro Arg Ser Pro Ala Val Ala Thr Glu	
1 5 10 15	
Ser Arg Arg Phe Ala Ala Ala Arg Trp Pro Gly Trp Arg Ser Leu Gln	
20 25 30	
Arg Pro Ala Arg Arg Ser Gly Arg Gly Gly Gly Ala Ala Pro Gly	
35 40 45	
Pro Tyr Pro Ser Ala Ala Pro Pro Pro Pro Gly Pro Gly Pro Pro Pro	
50 55 60	

Ser Arg Gln Ser Ser Pro Pro Ser Ala Ser Asp Cys Phe Gly Ser Asn  
65 70 75 80

Gly Asn Gly Gly Ala Phe Arg Pro Gly Ser Arg Arg Leu Leu Gly  
85 90 95

Leu Gly Gly Pro Pro Arg Pro Phe Val Val Val Leu Leu Pro Leu Ala  
100 105 110

Ser Pro Gly Ala Pro Pro Ala Ala Pro Thr Arg Ala Ser Pro Leu Gly  
115 120 125

Ala Arg Ala Ser Pro Pro Arg Ser Gly Val Ser Leu Ala Arg Pro Ala  
130 135 140

Pro Gly Cys Pro Arg Pro Ala Cys Glu Pro Val Tyr Gly Pro Leu Thr  
145 150 155 160

Met Ser Leu Lys Pro Gln  
165 170 175

Gln Gln Gln Gln Gln Gln Gln Gln Gln Pro Pro Pro Ala Ala  
180 185 190

Ala Asn Val Arg Lys Pro Gly Gly Ser Gly Leu Leu Ala Ser Pro Ala  
195 200 205

Ala Ala Pro Ser Pro Ser Ser Ser Val Ser Ser Ser Ala Thr  
210 215 220

Ala Pro Ser Ser Val Val Ala Ala Thr Ser Gly Gly Gly Arg Pro Gly  
225 230 235 240

Leu Gly Arg Gly Arg Asn Ser Asn Lys Gly Leu Pro Gln Ser Thr Ile  
245 250 255

Ser Phe Asp Gly Ile Tyr Ala Asn Met Arg Met Val His Ile Leu Thr  
260 265 270

Ser Val Val Gly Ser Lys Cys Glu Val Gln Val Lys Asn Gly Gly Ile  
275 280 285

Tyr Glu Gly Val Phe Lys Thr Tyr Ser Pro Lys Cys Asp Leu Val Leu  
290 295 300

Asp Ala Ala His Glu Lys Ser Thr Glu Ser Ser Ser Gly Pro Lys Arg  
305 310 315 320

Glu Glu Ile Met Glu Ser Ile Leu Phe Lys Cys Ser Asp Phe Val Val  
325 330 335

Val Gln Phe Lys Asp Met Asp Ser Ser Tyr Ala Lys Arg Asp Ala Phe  
340 345 350

Thr Asp Ser Ala Ile Ser Ala Lys Val Asn Gly Glu His Lys Glu Lys  
355 360 365

Asp Leu Glu Pro Trp Asp Ala Gly Glu Leu Thr Ala Asn Glu Glu Leu  
370 375 380

Glu Ala Leu Glu Asn Asp Val Ser Asn Gly Trp Asp Pro Asn Asp Met  
385 390 395 400

Phe Arg Tyr Asn Glu Glu Asn Tyr Gly Val Val Ser Thr Tyr Asp Ser  
405 410 415

Ser Leu Ser Ser Tyr Thr Val Pro Leu Glu Arg Asp Asn Ser Glu Glu  
420 425 430

Phe Leu Lys Arg Glu Ala Arg Ala Asn Gln Leu Ala Glu Glu Ile Glu  
435 440 445

Ser Ser Ala Gln Tyr Lys Ala Arg Val Ala Leu Glu Asn Asp Asp Arg  
450 455 460

Ser Glu Glu Glu Lys Tyr Thr Ala Val Gln Arg Asn Ser Ser Glu Arg  
465 470 475 480

Glu Gly His Ser Ile Asn Thr Arg Glu Asn Lys Tyr Ile Pro Pro Gly  
485 490 495

Gln Arg Asn Arg Glu Val Ile Ser Trp Gly Ser Gly Arg Gln Asn Ser  
500 505 510

Pro Arg Met Gly Gln Pro Gly Ser Gly Ser Met Pro Ser Arg Ser Thr  
515 520 525

Ser His Thr Ser Asp Phe Asn Pro Asn Ser Gly Ser Asp Gln Arg Val  
530 535 540

Val Asn Gly Gly Val Pro Trp Pro Ser Pro Cys Pro Ser Pro Ser Ser  
545 550 555 560

Arg Pro Pro Ser Arg Tyr Gln Ser Gly Pro Asn Ser Leu Pro Pro Arg  
565 570 575

Ala Ala Thr Pro Thr Arg Pro Pro Ser Arg Pro Pro Ser Arg Pro Ser  
580 585 590

Arg Pro Pro Ser His Pro Ser Ala His Gly Ser Pro Ala Pro Val Ser  
595 600 605

Thr Met Pro Lys Arg Met Ser Ser Glu Gly Pro Pro Arg Met Ser Pro  
610 615 620

Lys Ala Gln Arg His Pro Arg Asn His Arg Val Ser Ala Gly Arg Gly  
625 630 635 640

Ser Ile Ser Ser Gly Leu Glu Phe Val Ser His Asn Pro Pro Ser Glu  
645 650 655

Ala Ala Thr Pro Pro Val Ala Arg Thr Ser Pro Ser Gly Thr Trp  
660 665 670

Ser Ser Val Val Ser Gly Val Pro Arg Leu Ser Pro Lys Thr His Arg  
675 680 685

Pro Arg Ser Pro Arg Gln Asn Ser Ile Gly Asn Thr Pro Ser Gly Pro  
690 695 700

Val Leu Ala Ser Pro Gln Ala Gly Ile Ile Pro Thr Glu Ala Val Ala  
705 710 715 720

Met Pro Ile Pro Ala Ala Ser Pro Thr Pro Ala Ser Pro Ala Ser Asn  
725 730 735

Arg Ala Val Thr Pro Ser Ser Glu Ala Lys Asp Ser Arg Leu Gln Asp  
740 745 750

Gln Arg Gln Asn Ser Pro Ala Gly Asn Lys Glu Asn Ile Lys Pro Asn  
755 760 765

Glu Thr Ser Pro Ser Phe Ser Lys Ala Glu Asn Lys Gly Ile Ser Pro  
770 775 780

Val Val Ser Glu His Arg Lys Gln Ile Asp Asp Leu Lys Lys Phe Lys  
785 790 795 800

Asn Asp Phe Arg Leu Gln Pro Ser Ser Thr Ser Glu Ser Met Asp Gln  
805 810 815

Leu Leu Asn Lys Asn Arg Glu Gly Glu Lys Ser Arg Asp Leu Ile Lys  
820 825 830

Asp Lys Ile Glu Pro Ser Ala Lys Asp Ser Phe Ile Glu Asn Ser Ser  
835 840 845

Ser Asn Cys Thr Ser Gly Ser Ser Lys Pro Asn Ser Pro Ser Ile Ser  
850 855 860

Pro Ser Ile Leu Ser Asn Thr Glu His Lys Arg Gly Pro Glu Val Thr  
865 870 875 880

Ser Gln Gly Val Gln Thr Ser Ser Pro Ala Cys Lys Gln Glu Lys Asp  
885 890 895

Asp Lys Glu Glu Lys Lys Asp Ala Ala Glu Gln Val Arg Lys Ser Thr  
900 905 910

Leu Asn Pro Asn Ala Lys Glu Phe Asn Pro Arg Ser Phe Ser Gln Pro  
915 920 925

Lys Pro Ser Thr Thr Pro Thr Ser Pro Arg Pro Gln Ala Gln Pro Ser  
930 935 940

Pro Ser Met Val Gly His Gln Gln Pro Thr Pro Val Tyr Thr Gln Pro  
945 950 955 960

Val Cys Phe Ala Pro Asn Met Met Tyr Pro Val Pro Val Ser Pro Gly  
965 970 975

Val Gln Pro Leu Tyr Pro Ile Pro Met Thr Pro Met Pro Val Asn Gln  
980 985 990

Ala Lys Thr Tyr Arg Ala Val Pro Asn Met Pro Gln Gln Arg Gln Asp  
995 1000 1005

Gln His His Gln Ser Ala Met Met His Pro Ala Ser Ala Ala Gly Pro  
1010 1015 1020

Pro Ile Ala Ala Thr Pro Pro Ala Tyr Ser Thr Gln Tyr Val Ala Tyr  
1025 1030 1035 1040

Ser Pro Gln Gln Phe Pro Asn Gln Pro Leu Val Gln His Val Pro His  
1045 1050 1055

Tyr Gln Ser Gln His Pro His Val Tyr Ser Pro Val Ile Gln Gly Asn  
1060 1065 1070

Ala Arg Met Met Ala Pro Pro Thr His Ala Gln Pro Gly Leu Val Ser  
1075 1080 1085

Ser Ser Ala Thr Gln Tyr Gly Ala His Glu Gln Thr His Ala Met Tyr  
1090 1095 1100

Ala Cys Pro Lys Leu Pro Tyr Asn Lys Glu Thr Ser Pro Ser Phe Tyr  
1105 1110 1115 1120

Phe Ala Ile Ser Thr Gly Ser Leu Ala Gln Gln Tyr Ala His Pro Asn  
1125 1130 1135

Ala Thr Leu His Pro His Thr Pro His Pro Gln Pro Ser Ala Thr Pro  
1140 1145 1150

Thr Gly Gln Gln Gln Ser Gln His Gly Gly Ser His Pro Ala Pro Ser  
1155 1160 1165

Pro Val Gln His His Gln His Ala Ala Gln Ala Leu His Leu Ala  
1170 1175 1180

Ser Pro Gln Gln Gln Ser Ala Ile Tyr His Ala Gly Leu Ala Pro Thr  
1185 1190 1195 1200

Pro Pro Ser Met Thr Pro Ala Ser Asn Thr Gln Ser Pro Gln Asn Ser  
1205 1210 1215

Phe Pro Ala Ala Gln Gln Thr Val Phe Thr Ile His Pro Ser His Val  
1220 1225 1230

Gln Pro Ala Tyr Thr Asn Pro Pro His Met Ala His Val Pro Gln Ala  
1235 1240 1245

His Val Gln Ser Gly Met Val Pro Ser His Pro Thr Ala His Ala Pro  
1250 1255 1260

Met Met Leu Met Thr Thr Gln Pro Pro Gly Gly Pro Gln Ala Ala Leu  
1265 1270 1275 1280

Ala Gln Ser Ala Leu Gln Pro Ile Pro Val Ser Thr Thr Ala His Phe  
1285 1290 1295

Pro Tyr Met Thr His Pro Ser Val Gln Ala His His Gln Gln Leu  
1300 1305 1310

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1257 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: both

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2..1255

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

G CAC GAG GGG CCG CTC ACC ATG TCG CTG AAG CCG CAG CCG CAG CCG	46
His Glu Gly Pro Leu Thr Met Ser Leu Lys Pro Gln Pro Gln Pro	
1 5 10 15	
CCC GCG CCC GCC ACT GGC CGC AAG CCC GGC GGC CTG CTC TCG TCG	94
Pro Ala Pro Ala Thr Gly Arg Lys Pro Gly Gly Leu Leu Ser Ser	
20 25 30	
CCC GGC GCC GCG CCG GCC TCG GCC GCG GTG ACC TCG GCT TCC GTG GTG	142
Pro Gly Ala Ala Pro Ala Ser Ala Ala Val Thr Ser Ala Ser Val Val	
35 40 45	
CCG GCC CCG GCC GCG CCG GTG GCG TCT TCC TCG GCG GCC GCG GGC GGC	190
Pro Ala Pro Ala Ala Pro Val Ala Ser Ser Ala Ala Ala Gly Gly	
50 55 60	
GGG CGT CCC GGC CTG GGC AGA GGT CGG AAC AGT AGC AAA GGA CTG CCT	238
Gly Arg Pro Gly Leu Gly Arg Gly Arg Asn Ser Ser Lys Gly Leu Pro	
65 70 75	
CAG CCT ACG ATT TCT TTT GAT GGA ATC TAT GCA AAC GTG AGG ATG GTT	286
Gln Pro Thr Ile Ser Phe Asp Gly Ile Tyr Ala Asn Val Arg Met Val	
80 85 90 95	
CAT ATA CTT ACG TCA GTT GGA TCG AAA TGT GAA GTA CAA GTG AAA	334
His Ile Leu Thr Ser Val Val Gly Ser Lys Cys Glu Val Gln Val Lys	
100 105 110	
AAC GGA GGC ATA TAT GAA GGA GTT TTT AAA ACA TAC AGT CCT AAG TGT	382
Asn Gly Gly Ile Tyr Glu Gly Val Phe Lys Thr Tyr Ser Pro Lys Cys	
115 120 125	
GAC TTG GTA CTT GAT GCT GCA CAT GAG AAA AGT ACA GAA TCC AGT TCG	430
Asp Leu Val Leu Asp Ala Ala His Glu Lys Ser Thr Glu Ser Ser Ser	
130 135 140	
GGG CCA AAA CGT GAA GAA ATA ATG GAG AGT GTT TTG TTC AAA TGC TCA	478
Gly Pro Lys Arg Glu Glu Ile Met Glu Ser Val Leu Phe Lys Cys Ser	
145 150 155	
GAC TTC GTT GTG GTA CAG TTT AAA GAT ACA GAC TCC AGT TAT GCA CGG	526
Asp Phe Val Val Val Gln Phe Lys Asp Thr Asp Ser Ser Tyr Ala Arg	
160 165 170 175	
AGA GAT GCT TTT ACT GAC TCT GCT CTC AGC GCA AAG GTG AAT GGT GAG	574
Arg Asp Ala Phe Thr Asp Ser Ala Leu Ser Ala Lys Val Asn Gly Glu	
180 185 190	

CAC AAG GAG AAG GAC CTG GAG CCC TGG GAT GCA GGG GAG CTC ACG GCC	622
His Lys Glu Lys Asp Leu Glu Pro Trp Asp Ala Gly Glu Leu Thr Ala	
195 200 205	
AGC GAG GAG CTG GAG CTG GAG AAT GAT GTG TCT AAT GGA TGG GAC CCC	670
Ser Glu Glu Leu Glu Leu Glu Asn Asp Val Ser Asn Gly Trp Asp Pro	
210 215 220	
AAT GAC ATG TTT CGA TAT AAT GAA GAG AAT TAT GGT GTG GTG TCC ACA	718
Asn Asp Met Phe Arg Tyr Asn Glu Glu Asn Tyr Gly Val Val Ser Thr	
225 230 235	
TAT GAT AGC AGT TTA TCT TCA TAT ACG GTT CCT TTA GAA AGG GAC AAC	766
Tyr Asp Ser Ser Leu Ser Ser Tyr Thr Val Pro Leu Glu Arg Asp Asn	
240 245 250 255	
TCA GAA GAA TTT CTT AAA CGG GAG GCA AGG GCA AAC CAG TTA GCA GAA	814
Ser Glu Glu Phe Leu Lys Arg Glu Ala Arg Ala Asn Gln Leu Ala Glu	
260 265 270	
GAA ATT GAA TCC AGT GCT CAG TAC AAA GCT CGT GTC GCC CTT GAG AAT	862
Glu Ile Glu Ser Ser Ala Gln Tyr Lys Ala Arg Val Ala Leu Glu Asn	
275 280 285	
GAT GAC CGG AGT GAG GAA AAA TAC ACA GCA GTC CAG AGA AAC TGC	910
Asp Asp Arg Ser Glu Glu Lys Tyr Thr Ala Val Gln Arg Asn Cys	
290 295 300	
AGT GAC CGG GAG GGG CAT GGC CCC AAC ACT AGG GAC AAT AAA TAT ATT	958
Ser Asp Arg Glu Gly His Gly Pro Asn Thr Arg Asp Asn Lys Tyr Ile	
305 310 315	
CCT CCT GGA CAA AGA AAC AGA GAA GTC CTA TCC TGG GGA AGT GGG AGA	1006
Pro Pro Gly Gln Arg Asn Arg Glu Val Leu Ser Trp Gly Ser Gly Arg	
320 325 330 335	
CAG AGC TCA CCA CGG ATG GGC CAG CCT GGG CCA GGC TCC ATG CCG TCA	1054
Gln Ser Ser Pro Arg Met Gly Gln Pro Gly Pro Gly Ser Met Pro Ser	
340 345 350	
AGA GCT GCT TCT CAC ACT TCA GAT TTC AAC CCG AAC GCT GGC TCA GAC	1102
Arg Ala Ala Ser His Thr Ser Asp Phe Asn Pro Asn Ala Gly Ser Asp	
355 360 365	
CAA AGA GTA GTT AAT GGA GGT GTT CCC TGG CCA TCG CCT TGC CCA TCT	1150
Gln Arg Val Val Asn Gly Gly Val Pro Trp Pro Ser Pro Cys Pro Ser	
370 375 380	
CCT TCC TCT CGC CCA CCT TCT CGC TAC CAG TCA GGT CCC AAC TCT CTT	1198
Pro Ser Ser Arg Pro Pro Ser Arg Tyr Gln Ser Gly Pro Asn Ser Leu	
385 390 395	

CCA CCT CGG GCA GCC ACC CCT ACA CGG CCT CGT GCC GAA TTC CTG CAG 1246  
Pro Pro Arg Ala Ala Thr Pro Thr Arg Pro Arg Ala Glu Phe Leu Gln  
400 405 410 415

CCC GGG GAT CC 1257  
Pro Gly Asp

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 418 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

His Glu Gly Pro Leu Thr Met Ser Leu Lys Pro Gln Pro Gln Pro Pro  
1 5 10 15

Ala Pro Ala Thr Gly Arg Lys Pro Gly Gly Leu Leu Ser Ser Pro  
20 25 30

Gly Ala Ala Pro Ala Ser Ala Ala Val Thr Ser Ala Ser Val Val Pro  
35 40 45

Ala Pro Ala Ala Pro Val Ala Ser Ser Ala Ala Gly Gly Gly  
50 55 60

Arg Pro Gly Leu Gly Arg Gly Arg Asn Ser Ser Lys Gly Leu Pro Gln  
65 70 75 80

Pro Thr Ile Ser Phe Asp Gly Ile Tyr Ala Asn Val Arg Met Val His  
85 90 95

Ile Leu Thr Ser Val Val Gly Ser Lys Cys Glu Val Gln Val Lys Asn  
100 105 110

Gly Gly Ile Tyr Glu Gly Val Phe Lys Thr Tyr Ser Pro Lys Cys Asp  
115 120 125

Leu Val Leu Asp Ala Ala His Glu Lys Ser Thr Glu Ser Ser Ser Gly  
130 135 140

Pro Lys Arg Glu Glu Ile Met Glu Ser Val Leu Phe Lys Cys Ser Asp  
145 150 155 160

Phe Val Val Val Gln Phe Lys Asp Thr Asp Ser Ser Tyr Ala Arg Arg  
165 170 175

Asp Ala Phe Thr Asp Ser Ala Leu Ser Ala Lys Val Asn Gly Glu His  
180 185 190

Lys Glu Lys Asp Leu Glu Pro Trp Asp Ala Gly Glu Leu Thr Ala Ser  
195 200 205

Glu Glu Leu Glu Leu Glu Asn Asp Val Ser Asn Gly Trp Asp Pro Asn  
210 215 220

Asp Met Phe Arg Tyr Asn Glu Glu Asn Tyr Gly Val Val Ser Thr Tyr  
225 230 235 240

Asp Ser Ser Leu Ser Ser Tyr Thr Val Pro Leu Glu Arg Asp Asn Ser  
245 250 255

Glu Glu Phe Leu Lys Arg Glu Ala Arg Ala Asn Gln Leu Ala Glu Glu  
260 265 270

Ile Glu Ser Ser Ala Gln Tyr Lys Ala Arg Val Ala Leu Glu Asn Asp  
275 280 285

Asp Arg Ser Glu Glu Glu Lys Tyr Thr Ala Val Gln Arg Asn Cys Ser  
290 295 300

Asp Arg Glu Gly His Gly Pro Asn Thr Arg Asp Asn Lys Tyr Ile Pro  
305 310 315 320

Pro Gly Gln Arg Asn Arg Glu Val Leu Ser Trp Gly Ser Gly Arg Gln  
325 330 335

Ser Ser Pro Arg Met Gly Gln Pro Gly Pro Gly Ser Met Pro Ser Arg  
340 345 350

Ala Ala Ser His Thr Ser Asp Phe Asn Pro Asn Ala Gly Ser Asp Gln  
355 360 365

Arg Val Val Asn Gly Gly Val Pro Trp Pro Ser Pro Cys Pro Ser Pro  
370 375 380

Ser Ser Arg Pro Pro Ser Arg Tyr Gln Ser Gly Pro Asn Ser Leu Pro  
385 390 395 400

Pro Arg Ala Ala Thr Pro Thr Arg Pro Arg Ala Glu Phe Leu Gln Pro  
405 410 415

Gly Asp

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GGGCCCCCTCA CCATGTCG

18

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CGGGCTTGCG GACATTGG

18

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CCGCGGGCTGC CAATGTCC

18

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GTAACCGTTC GGCGCCCG

18

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GGCTCCCGGC GGCTCCTT

18

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TGCTGCTGCT GCTGGGGCTT CAG

23

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CCGCCCGCTC CTCACGTGT

19

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

ACCCCCGAGA AAGCAACC

18

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CCGTTGCCGT TGCTACCA

18

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

TTCTCATGTG CGGCATCAAG

20

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 346 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Val Tyr Gly Pro Leu Thr Met Ser Leu Lys Pro Gln Gln Gln Gln  
1 5 10 15

Gln  
20 25 30

Gln Pro Pro Pro Ala Ala Ala Asn Val Arg Lys Pro Gly Gly Ser Gly  
35 40 45

Leu Leu Ala Ser Pro Ala Ala Pro Ser Pro Ser Ser Ser Val  
50 55 60

Ser Ser Ser Ala Thr Ala Pro Ser Ser Val Val Ala Ala Thr Ser  
65 70 75 80

Gly Gly Gly Arg Pro Gly Leu Gly Arg Gly Arg Asn Ser Asn Lys Gly  
85 90 95

Leu Pro Gln Ser Thr Ile Ser Phe Asp Gly Ile Tyr Ala Asn Met Arg  
100 105 110

Met Val His Ile Leu Thr Ser Val Val Gly Ser Lys Cys Glu Val Gln  
 115 120 125  
 Val Lys Asn Gly Gly Ile Tyr Glu Gly Val Phe Lys Thr Tyr Ser Pro  
 130 135 140  
 Lys Cys Asp Leu Val Leu Asp Ala Ala His Glu Lys Ser Thr Glu Ser  
 145 150 155 160  
 Ser Ser Gly Pro Lys Arg Glu Glu Ile Met Glu Ser Ile Leu Phe Lys  
 165 170 175  
 Cys Ser Asp Phe Val Val Val Gln Phe Lys Asp Met Asp Ser Ser Tyr  
 180 185 190  
 Ala Lys Arg Asp Ala Phe Thr Asp Ser Ala Ile Ser Ala Lys Val Asn  
 195 200 205  
 Gly Glu His Lys Glu Lys Asp Leu Glu Pro Trp Asp Ala Gly Glu Leu  
 210 215 220  
 Thr Ala Asn Glu Glu Leu Glu Ala Leu Glu Asn Asp Val Ser Asn Gly  
 225 230 235 240  
 Trp Asp Pro Asn Asp Met Phe Arg Tyr Asn Glu Glu Asn Tyr Gly Val  
 245 250 255  
 Val Ser Thr Tyr Asp Ser Ser Leu Ser Ser Tyr Thr Val Pro Leu Glu  
 260 265 270  
 Arg Asp Asn Ser Glu Glu Phe Leu Lys Arg Glu Ala Arg Ala Asn Gln  
 275 280 285  
 Leu Ala Glu Glu Ile Glu Ser Ser Ala Gln Tyr Lys Ala Arg Val Ala  
 290 295 300  
 Leu Glu Asn Asp Asp Arg Ser Glu Glu Glu Lys Tyr Thr Ala Val Gln  
 305 310 315 320  
 Arg Asn Ser Ser Glu Arg Glu Gly His Ser Ile Asn Thr Arg Glu Asn  
 325 330 335  
 Lys Tyr Ile Pro Pro Gly Gln Arg Asn Arg  
 340 345

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 326 amino acids
- (B) TYPE: amino acid

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

His Glu Gly Pro Leu Thr Met Ser Leu Lys Pro Gln Pro Gln Pro Pro  
1 5 10 15

Ala Pro Ala Thr Gly Arg Lys Pro Gly Gly Gly Leu Leu Ser Ser Pro  
20 25 30

Gly Ala Ala Pro Ala Ser Ala Ala Val Thr Ser Ala Ser Val Val Pro  
35 40 45

Ala Pro Ala Ala Pro Val Ala Ser Ser Ser Ala Ala Ala Gly Gly Gly  
50 55 60

Arg Pro Gly Leu Gly Arg Gly Arg Asn Ser Ser Lys Gly Leu Pro Gln  
65 70 75 80

Pro Thr Ile Ser Phe Asp Gly Ile Tyr Ala Asn Val Arg Met Val His  
85 90 95

Ile Leu Thr Ser Val Val Gly Ser Lys Cys Glu Val Gln Val Lys Asn  
100 105 110

Gly Gly Ile Tyr Glu Gly Val Phe Lys Thr Tyr Ser Pro Lys Cys Asp  
115 120 125

Leu Val Leu Asp Ala Ala His Glu Lys Ser Thr Glu Ser Ser Ser Gly  
130 135 140

Pro Lys Arg Glu Glu Ile Met Glu Ser Val Leu Phe Lys Cys Ser Asp  
145 150 155 160

Phe Val Val Val Gln Phe Lys Asp Thr Asp Ser Ser Tyr Ala Arg Arg  
165 170 175

Asp Ala Phe Thr Asp Ser Ala Leu Ser Ala Lys Val Asn Gly Glu His  
180 185 190

Lys Glu Lys Asp Leu Glu Pro Trp Asp Ala Gly Glu Leu Thr Ala Ser  
195 200 205

Glu Glu Leu Glu Leu Glu Asn Asp Val Ser Asn Gly Trp Asp Pro Asn  
210 215 220

Asp Met Phe Arg Tyr Asn Glu Glu Asn Tyr Gly Val Val Ser Thr Tyr  
225 230 235 240

Asp Ser Ser Leu Ser Ser Tyr Thr Val Pro Leu Glu Arg Asp Asn Ser  
245 250 255

Glu Glu Phe Leu Lys Arg Glu Ala Arg Ala Asn Gln Leu Ala Glu Glu  
260 265 270

Ile Glu Ser Ser Ala Gln Tyr Lys Ala Arg Val Ala Leu Glu Asn Asp  
275 280 285

Asp Arg Ser Glu Glu Glu Lys Tyr Thr Ala Val Gln Arg Asn Cys Ser  
290 295 300

Asp Arg Glu Gly His Gly Pro Asn Thr Arg Asp Asn Lys Tyr Ile Pro  
305 310 315 320

Pro Gly Gln Arg Asn Arg  
325

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 275 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Leu Ala Pro Gln Pro Pro Pro Gln Gln His Gln Glu Arg Pro Gly  
1 5 10 15

Ala Ala Ala Ile Gly Ser Ala Arg Gly Gln Ser Thr Gly Lys Gly Pro  
20 25 30

Pro Gln Ser Pro Val Phe Glu Gly Val Tyr Asn Asn Ser Arg Met Leu  
35 40 45

His Phe Leu Thr Ala Val Val Gly Ser Thr Cys Asp Val Lys Val Lys  
50 55 60

Asn	Gly	Thr	Thr	Tyr	Glu	Gly	Ile	Phe	Lys	Thr	Leu	Ser	Ser	Lys	Phe	
65					70					75				80		
Glu Leu Ala Val Asp Ala Val His Arg Lys Ala Ser Glu Pro Ala Gly																
					85					90				95		
Gly Pro Arg Arg Glu Asp Ile Val Asp Thr Met Val Phe Lys Pro Ser																
					100					105				110		
Asp Val Met Leu Val His Phe Arg Asn Val Asp Phe Asn Tyr Ala Thr																
					115					120				125		
Lys Asp Lys Phe Thr Asp Ser Ala Ile Ala Met Asn Ser Lys Val Asn																
					130					135				140		
Gly Glu His Lys Glu Lys Val Leu Gln Arg Trp Glu Gly Gly Asp Ser																
					145					150				155		160
Asn Ser Asp Asp Tyr Asp Leu Glu Ser Asp Met Ser Asn Gly Trp Asp																
					165					170				175		
Pro Asn Glu Met Phe Lys Phe Asn Glu Glu Asn Tyr Gly Val Lys Thr																
					180					185				190		
Thr Tyr Asp Ser Ser Leu Ser Ser Tyr Thr Val Pro Leu Glu Lys Asp																
					195					200				205		
Asn Ser Glu Glu Phe Arg Gln Arg Glu Leu Arg Ala Ala Gln Leu Ala																
					210					215				220		
Arg Glu Ile Glu Ser Ser Pro Gln Tyr Arg Leu Arg Ile Ala Met Glu																
					225					230				235		240
Asn Asp Asp Gly Arg Thr Glu Glu Lys His Ser Ala Val Gln Arg																
					245					250				255		
Gln Gly Ser Gly Arg Glu Ser Pro Ser Leu Ala Ser Arg Glu Gly Lys																
					260					265				270		
Tyr Ile Pro																
					275											

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CCAATGTCCG CAAGCCCC

18

Cl  
cont